# Data, Determination, and Decisions: A Branch Campus Project for HLC Persistence and **Completion Academy**

### **Project Origins**

The University project began in October 2013 Divided into 4 teams to focus on different student populations:

- Macomb Team freshmen to sophomore retention
- Quad Cities Team retention of transfer students
- Distance Learning Team retention of online students
- Data Team creating the tools needed to gather and understand patterns of retention

### **Project Relationship to** University initiative

- Similar goals, processes, and tools
- Identify and define targeted student population
- Discover retention patterns among targeted group
- Detect factors that inhibit student persistence
- Select student attributes for data warehouse development and analysis
- Communicate progress university-wide

### Institutional Characteristics

### Western Illinois University:

- Public State Supported University
- Total Enrollment: 11,094
- Locations in Macomb (main campus) and Moline (branch campus)

### WIU Quad Cities Campus:

- Urban, non-residential campus
- Total Enrollment: 1,531
- Top majors: Engineering, Management, Accountancy, Elementary Education
- 40% of undergraduates are part-time
- 19% minority enrollment
- 58% female enrollment

- Used data warehouse to gather information on undergraduate transfer students from 2011-14
- N = 947
- 14 attributes including gender, ethnicity, transfer GPA, class level, enrollment status, application submission period, home state and academic status after first term among others
- Used Weka (machine learning) to build a model visualizing data; Classified data with J48 (C4.5) algorithm with 10-fold cross-validation, and Information Gain Attribute Evaluator
- Weka correctly classified 82.7% instances
- Used attribute evaluator to cross-validate decision tree results •
- Plan to re-run J48 using only the top n attributes (number to be determined by team)



Audrey Adamson & Debbie Kepple-Mamros Western Illinois University – Quad Cities, Moline, IL

# Methodology

The goal was to allow the data to show us what was most important rather than start with assumptions that we test for statistical significance.

### **Results – Preliminary Decision Tree**



## **Cross-Validation** Attempts

Ranked Attributes using Weka's **Information Gain Attribute Evaluator** 

1.Enrollment Status (0.288724)

2.Academic Status After First Semester (0.12733)

3.Application Submission Period (0.037529)

4.State of Residency (0.014678)

### **Conclusions to Date**

•Changed our original focus on "linkages" (dually enrolled) students due to small sample

•Weka is a useful tool but not yet widely understood or trusted. Need to educate and improve communication.

•Our own use of the data warehouse includes a steep learning curve.

•The data warehouse is a work in progress, as are agreed-upon definitions terminology.

•Our preliminary decision tree produced results that confirmed assumptions.

•Cross-validation sheds light on additional important attributes.



WESTERN ILLINOIS UNIVERSITY **Quad** Cities